BEBR
FACULTY WORKING PAPER NO. 90-1659

On the Generic Nonconvergence of Bayesian Actions and Beliefs

Mark Feldman

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 90-1659

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

June 1990

On the Generic Nonconvergence of Bayesian
Actions and Beliefs

Mark Feldman

Department of Economics
University of Illinois at Urbana-Champaign

I have benefited from opportunities to discuss this research with Andrew Barron, Larry Blume, Roger Koenker, and Andrew McLennan. I am also grateful for the comments of seminar participants at the University of Minnesota and at the conference on "Endogenous Learning in Economic Environments," Cornell University, June 1988. Of course, only I am responsible for any errors.

ABSTRACT

Suppose Y_n is a sequence of i.i.d. random variables taking values in Y, a complete, separable, non-finite metric space. The probability law indexed by $\theta \in \Theta$, is unknown to a Bayesian statistician with prior μ, observing this process. Generalizing Freedman [1965, *Annals of Mathematical Statistics*], we show that "generically" (i.e., for a residual family of (θ, μ) pairs) the posterior beliefs do not weakly converge to a point-mass at the "true" θ. Furthermore, for every open set $G \subset \Theta$, generically, the Bayesian will attach probability arbitrarily close to one to G infinitely often.

The above result is applied to a two-armed bandit problem with geometric discounting where arm k yields an outcome in a complete, separable metric space Y_k. If the infimum of the possible rewards from playing arm k is less than the infimum from playing arm k′, then arm k is (generically) chosen only finitely often. If the infimum of the rewards are equal, then both arms are played infinitely often.

JEL Classification Numbers: 022, 026, 211.

1. INTRODUCTION

There has been a flurry of interest in studying the asymptotic dynamics of Bayesian learning and control in economic environments. In one set of papers including Easley and Kiefer (1988), Easley and Kiefer (1989), Kiefer and Nyarko (1988), Kiefer and Nyarko (1989), McLennan (1987), Feldman and McLennan (1989), and Bikhchandani and Sharma (1990), the authors analyze single agent decision problems in which there is a tradeoff between current period expected reward and the expected value of the information generated by the current period action. In another strand of the literature, Blume and Easley (1984), Bray and Kreps (1987), Feldman (1987a), and Feldman (1987b) focus on properties of the tail of the sequence of beliefs and outcomes for economies with many passively learning agents. Specifically, these latter papers consider whether Bayesian learning by agents with a correct specification of the underlying structure but uncertainty regarding the parameter values is a sufficient condition to assure convergence to a stationary rational expectations equilibrium.

These articles, as well as important earlier contributions of Cyert and DeGroot (1973), Rothschild (1974), and Townsend (1978), have the following common framework. From the vantage point of the economic actors, the set of possible complete descriptions of the relevant time-invariant economic data can be represented as a separable metric space Θ with Borel σ-field $B(\Theta)$. The actors in the model, uncertain as to the "true" $\theta_0 \in \Theta$, have prior beliefs μ on $(\Theta, B(\Theta))$[1], and an induced probability P_μ on an infinite horizon outcome space. Denote by $\{\mu_t\}$ the sequence of posterior beliefs. A result common to this literature is a "theorem" that with probability one $\mu_t \Rightarrow \mu_\infty$, where μ_∞ is the posterior probabililty conditioned on the limit sub-σ-field.

In most of the recent papers (exceptions are Feldman and McLennan (1989) and Bikhchandani and Sharma (1990)), this a.s. convergence is established by using the fact that the the sequence of posterior beliefs are a martingale with respect to the probability P_μ. It follows from the Martingale Convergence Theorem that with P_μ probability one, the Bayesian beliefs converge to some (possibly random) limit belief.. In contrast, consider the distribution of outcomes and beliefs with respect to the probability measure P_{θ_0}, the probability induced by the "true" parameter θ_0. Intuitively, P_{θ_0} is the belief of a passive observer who attaches probability one to θ_0 being the truth. One might also inquire as to whether convergence of Bayesian beliefs is obtained with respect to the measure P_{θ_0}. A major point of this paper is to stress that without additional conditions, the answer is negative.

To elaborate on this distinction, suppose that in period t = 0, 1, ... , agents observe outcomes in a separable metric space Y. Given the behavioral rules of the agents, each parameter value $\theta \in \Theta$ induces a probability measure P_θ on the product space Y^∞. The prior μ induces a measure P_μ on Y^∞ defined by $P_\mu(A) = \int P_\theta(A) P_\mu(d\theta)$. The application of the Martingale Convergence Theorem yields a.s. convergence of $\{\mu_t\}$, where the a.s. statement is with respect to the probability measure P_μ. But this does not imply that for any particular θ that $\mu_t \Rightarrow \mu_\infty$ with P_θ probability one, even if θ is in the support of μ.

One might hope to establish a result that for a "large" class of priors, posteriors converge for a "large" class of parameter values. When Θ can be embedded in finite-dimensional Euclidean space, one has recourse to Lebesgue measure m (restricted to Θ) as a natural notion of size. Then if $m << \mu$ the exceptional θ set $\{\theta\colon P_\theta(\{\mu_t \not\Rightarrow \mu\}) < 1\}$,

has m measure zero. But in many naturally occurring settings Θ is not finite dimensional, and since there is no infinite dimensional analogue of Lebesgue measure, a measure-theoretic criterion is unavailable.

In lieu of a reference measure to evaluate size, the customary procedure is to resort to the topological notion of category. Residual subsets are deemed to be large or *generic*, and subsets of first category (which are complements of residual subsets) are regarded as small. Freedman (1965) proved that when outcomes are I.I.D. taking values in a countable set, that for a residual set of parameter values and priors, posterior beliefs do not converge. In Section 3 of this paper we extend Freedman's result to outcomes in non-finite, complete, separable metric spaces.

Using the results of Section 3, in Sections 4 and 5 of the paper we analyze a two-armed bandit problem with geometric discounting where arm k yields an outcome in a complete, separable metric space Y_k. If the infimum of the possible rewards from playing arm k is less than the infimum from playing arm k′, then for a residual family of parameter values and priors, arm k is with P_θ probabilty one chosen only finitely often. If the infimum of the rewards are equal, then both arms are played infinitely often.

2. NOTATION AND MATHEMATICAL PRELIMINARIES

2.1. *Notational Conventions and Definitions*

The set of real numbers is denoted by **R**. If X is a topological space, then the Borel σ-field is denoted by $B(X)$. The set of probability measures on $(X, B(X))$ is denoted by $P(X)$. For $x \in X$, the Dirac measure $\delta_x \in P(X)$ is defined by $\delta_x(A) = 1$ if $x \in A$.

If (X, d) is a metric space, $f: X \to \mathbf{R}$ is a Lipschitz function if for some $K < \infty$, $\sup_{x \neq y}\{|f(x) - f(y)|/d(x,y)\} < K$. If f is Lipschitz, the Lipschitz seminorm $\|f\|_L$ is defined $\|f\|_L = \sup_{x \neq y}\{|f(x) - f(y)|/d(x,y)\}$. If f is a bounded Lipschitz function, the *bounded Lipschitz* norm is $\|f\|_{BL} = \|f\|_L + \|f\|_\infty$ where $\|f\|_\infty$ denotes the usual sup norm. The set of all real-valued, bounded Lipschitz functions on (X, d) is denoted by BL(X, d). Endowed with the bounded Lipschitz norm, BL(X, d) is a Banach space (see e.g. Dudley (1989, Section 11.2)).

The *dual bounded Lipschitz* or *Dudley* metric β on $P(X)$ is defined by

$$\beta(P, Q) = \sup\{|\int f\, dP - \int f\, dQ| : \|f\|_{BL} \leq 1\},$$

for $P, Q \in P(X)$. If X is separable, β metrizes the topology of weak convergence on $P(X)$. Further details on the properties of β can be found in Dudley (1966) and Dudley (1989).

2.2. *A Brief Review of Baire Category Theory*

For ease of reference, we summarize some needed facts pertaining to Baire category. Standard references include Kelley (1985, pp. 200-203), Oxtoby (1980) and Royden (1988, Section 7.8). Let X be a metric space. A set $E \subset X$ is *nowhere dense* if $\bar{E}$ has empty interior. A set E is of *first category* or *meager* if it is the union of a countable collection of nowhere dense sets. If a set is not of first category then it is of *second category*. The complement of a set of first category is a *residual set*.

According to the Theorem of Baire Royden (1988, Theorem 7.27), if X is a complete metric space then the intersection of a countable family of open dense subsets of X is itself a dense subset of X.

3. GENERIC NONCONVERGENCE OF POSTERIORS WITH I.I.D. OUTCOMES IN COMPLETE, SEPARABLE METRIC SPACES

3.1. *Assumptions and Results*

We first describe an index set Λ and a sequence $Z_1, Z_2, \ldots$ of i.i.d. random variables defined on a probability space (Σ, S, P_λ) where $\lambda \in \Lambda$. The natural interpretation will be that the outcomes are sequentially observed by a Bayesian statistician for whom the "true" λ is initially unknown and has prior belief μ. Building upon the work of Freedman (1965), we will investigate the topological size of the set of pairs (λ, μ) for which the sequence of Bayesian posterior beliefs converges P_λ a.s. or in P_λ probability to some limit posterior belief.

The sequence $\{Z_n\}$ takes values in Z, a non-finite, complete, separable metric space with Borel σ-field $B(Z)$. The probability distribution of Z_n is an element of $\Lambda = \{\lambda \in P(Z): \lambda << v\}$, where v is a σ-finite measure on $(Z, B(Z))$ with non-finite support. Without loss of generality we work in representation space and so define $\Sigma = Z^\infty$, $S = B(Z) \times B(Z) \times \ldots$, and $P_\lambda = \lambda \times \lambda \times \ldots$. The function $Z_n: \Sigma \to Z$ is the projection of Σ onto Z, defined by $Z_n(z_1, z_2, \ldots) = z_n$.

To address the question of convergence of posterior beliefs we need topologies on Λ and $P(\Lambda)$. We will make use of two topologies on Λ, the total variation topology $\mathcal{T}_1$ and the topology of weak convergence $\mathcal{T}_w$. $\mathcal{T}_1$ is induced by the L^1 metric d_1 defined by $d_1(\lambda, \lambda') = \int_Z \left|\frac{d\lambda}{dv} - \frac{d\lambda'}{dv}\right| dv$. An essential fact is that (Λ, d_1) is a complete, separable metric space. (Completeness follows from the completeness of $L^1(Z, B(Z), v)$, and for separability see e.g., Strasser (1985, Lemma 4.1).) In contrast, defining β_Λ as the Dudley metric on Λ (which generates the topology $\mathcal{T}_w$), the metric space (Λ, β_Λ) is separable, but not complete. Conveniently, the Borel σ-field of $(\Lambda, \mathcal{T}_1)$ is the same as the Borel σ-field of $(\Lambda, \mathcal{T}_w)$ (see Strasser (1985, Theorem 4.7). So without ambiguity we can denote the Borel sets of Λ by $B(\Lambda)$.

A *prior distribution* is a probability measure μ on $(\Lambda, B(\Lambda))$. As indicated above, informally one can imagine that there is a Bayesian statistician who may not know the "true" λ, but has a prior μ. Since Λ has two topologies, there are two weak topologies on $P(\Lambda)$ denoted in the obvious way by $\mathcal{W}_1$ and $\mathcal{W}_w$, with $\mathcal{W}_w$ weaker than $\mathcal{W}_1$. $\mathcal{W}_w$ and $\mathcal{W}_1$ generate the same σ-field, which we denote by $B(P(\Lambda))$. Convergence with respect to the $\mathcal{W}_1$ topology is denoted by $\Rightarrow$. Convergence with respect to the $\mathcal{W}_w$ topology is denoted by $\stackrel{w}{\Rightarrow}$. In this section of the paper the symbol β denotes the Dudley metric on $P(\Lambda)$ with respect to the d_1 metric on Λ.. It follows from Billingsley (1968, p. 239) and Dudley (1989, Corollary 11.5.5)) that the metric space $(P(\Lambda), \beta)$ is complete and separable with β generating the topology $\mathcal{W}_1$.

The *updating rule* $\Gamma: P(\Lambda) \times Z \to P(\Lambda)$ is a measurable function with the property that for each $\mu \in P(\Lambda)$, $\Gamma(\mu, \cdot)$ is a regular version of conditional probability with respect to the prior probability μ. The existence of such a

function is established by Dynkin and Yushkevich (1979, p. 263). The n-period updating rule is Γ_n: $P(\Lambda) \times \Sigma \to P(\Lambda)$, recursively defined by $\Gamma_1(\mu, \sigma) = \Gamma(\mu, Z_1(\sigma))$ and $\Gamma_n(\mu, \sigma) = \Gamma(\Gamma_{n-1}(\mu, \sigma), Z_n(\sigma))$ for $n \geq 2$.

A pair $(\lambda, \mu) \in \Lambda \times P(\Lambda)$ is *$\mathscr{T}_1$-consistent* if $P_\lambda(\{\sigma: \Gamma_n(\mu, \sigma) \Rightarrow \delta_\lambda\}) = 1$. A pair $(\lambda, \mu) \in \Lambda \times P(\Lambda)$ is *$\mathscr{T}_w$-consistent* if $P_\lambda(\{\sigma: \Gamma_n(\mu, \sigma) \stackrel{w}{\Rightarrow} \delta_\lambda\}) = 1$. (When v has countable support, $\mathscr{T}_1 = \mathscr{T}_w$ and so the two definitions of consistency are identical.) Freedman (1963) proved that when Z is finite and μ has full support, that (λ, μ) is consistent for all $\lambda \in P(\Lambda)$. It would be natural to conjecture a similar result for when Z is not finite. Indeed, the consistency result for the finite outcome case has been generalized in a well-known paper of Schwartz (1965) and more recently by Barron (1988). However, Freedman (1965) demonstrated that in a topolological sense "most" pairs are not consistent (in either sense) when v has countable support, even if is required that the prior μ has full support. More precisely, defining $S = \{\mu \in P(\Lambda): \text{supp}\, \mu = \Lambda\}$, Freedman proved the striking result that there exist sets $R_\Lambda \subset \Lambda$ and $R_{P(\Lambda)} \subset S$, residual in Λ and $P(\Lambda)$ respectively (which implies that $R_\Lambda \times R_{P(\Lambda)}$ is residual in $\Lambda \times P(\Lambda)$), such that for $(\lambda, \mu) \in R_\Lambda \times R_{P(\Lambda)}$:

$\limsup_{n\to\infty} \int_\Sigma \Gamma_n(\mu, \sigma)(G)\, P_\lambda(d\sigma) = 1$, for all nonempty open subsets $G \subset \Lambda$. A corollary is that for (λ, μ) in the residual set $R_\Lambda \times R_{P(\Lambda)}$, $P_\lambda(\{\sigma: \Gamma_n(\mu, \sigma) \Rightarrow \delta_\lambda\}) = 0$.

In the next subsection we show that these non-convergence results of Freedman (1965) extend to the case where v is any σ-finite measure with non-finite support. Endowing Λ with the d_1 metric, and defining $\mathscr{R} = \{(\lambda, \mu) \in \Lambda \times P(\Lambda): \limsup_{n\to\infty} \int_\Sigma \Gamma_n(\mu, \sigma)(G)\, P_\lambda(d\sigma) = 1 \text{ for all open } G \subset \Lambda\}$, we prove the following theorem and corollary.

THEOREM 3.10. $\mathscr{R}$ is a residual subset of $\Lambda \times P(\Lambda)$. Furthermore, defining $S = \{\mu \in P(\Lambda): \text{supp}\, \mu = \Lambda\}$ and $\mathscr{R}_S \doteq \mathscr{R} \cap (\Lambda \times S)$, $\mathscr{R}_S$ is a residual subset of $\Lambda \times P(\Lambda)$.

COROLLARY 3.11. All $(\lambda, \mu) \in \mathscr{R}$ are neither $\mathscr{T}_1$ or $\mathscr{T}_w$-consistent.

3.2. *Proof of Theorem and Corollary*

While the basic structure of the proof is closely resembles the proof of Freedman (1965) of his Theorem, some modification and extension is required to adapt the argument to cover a non-discrete outcome space. In particular, Proposition 3.5 requires a different method of proof than the comparable intermediate result in Freedman (1965).

We start with some definitions. Unless otherwise indicated, where relevant it should be understood that Λ is endowed with the $\mathscr{T}_1$ topology. Define h: $\Lambda \times Z \to \mathbf{R}$ as $(B(\Lambda) \times B(Z))$ measurable function such that $h(\lambda, \cdot)$ is a density for λ with respect to v (a proof of existence of such a function is provided by Strasser (1985, Lemma 4.6)). Let $\Lambda_+ = \{\lambda \in \Lambda: h(\lambda, z) > 0, \text{ v a.e.}\}$ and define $\Lambda_0 = \sim\Lambda_+$. The set of probability measures on $(\Lambda, B(\Lambda))$ that assign strictly positive probability to Λ_+ is $P_+(\Lambda) = \{\mu \in P(\Lambda): \mu(\Lambda_+) > 0\}$.

Λ_+ and $P_+(\Lambda)$ are topologically "large" in the sense that each is a residual subset of a complete, separable metric space. In contrast, Λ_0 and $P(\Lambda_0) = \{\mu \in P(\Lambda): \mu(\Lambda_0) = 1\}$ are of first category, albeit dense in respectively Λ and $P(\Lambda)$.

LEMMA 3.1. *$\Lambda_+ \subset \Lambda$ and $P_+(\Lambda) \subset P(\Lambda)$ are respectively dense* G_δ *(and hence residual) subsets of* Λ *and* $P_+(\Lambda)$. $\Lambda_0 \subset \Lambda$ *and* $P(\Lambda_0) \subset P(\Lambda)$ *are dense sets of first category.*

Proof. (i) We first establish the properties of Λ_+ and $P(\Lambda_+)$. Define $\Lambda^j = \{\lambda \in \Lambda: v(\{z: h(\lambda, z) = 0\}) < j^{-1}\}$ for $j = 1, 2, \ldots$. Λ^j is an open, dense subset of Λ. So by the Baire Category Theorem (see e.g., Royden (1988, Theorem 7.27)) $\Lambda_+ = \cap_{j=1}^{\infty} \Lambda^j$ is a dense G_δ set and hence residual (Royden (1988, Theorem 7.30)).

The claim that $P_+(\Lambda)$ is a dense G_δ, follows from Theorem 3.15 of Dubins and Freedman (1964). (Dubins and Freedman have a compactness assumption in Section 3 of their paper; but inspection of their proof reveals that completeness and separability is a sufficient condition.)

(ii) We now verify the properties of Λ_0 and $P(\Lambda_0)$. Since $\Lambda_0 = {\sim}\Lambda_+$ and $P(\Lambda_0) = {\sim}P_+(\Lambda)$, and Λ_+ and $P_+(\Lambda)$ are residual, by definition Λ_0 and $P(\Lambda_0)$ are of first category. To prove denseness, for arbitrary $\lambda \in \Lambda$ choose a sequence $\lambda^k \to \lambda$ such that $h(\lambda^k, z) = 0$ on a set of positive v measure and $h(\lambda^k, \cdot)$ converges in v measure to $h(\lambda, \cdot)$. (The existence of such a sequence follows from the fact that for every $\alpha \in [0, 1]$ there exists $A_\alpha \in B(Z)$ such that $\alpha = \int_{A_\alpha} h(\lambda, z)\, v(dz)$.) But then $\lambda^k \to \lambda$, establishing the denseness of Λ_0. The density of $P(\Lambda_0)$ now follows from Theorem II.6.3 of Parthasarathy (1967). ■

Let $D = \{\alpha_1, \alpha_2, \ldots\}$ be a countable, dense subset of Λ_+ and hence dense subset of Λ. We now construct a sequence $M_1, M_2, \ldots$ with $M_k \subset P(\Lambda)$ such that for all $\lambda \in \Lambda_+$ and $\mu \in M_k$, $\Gamma_n(\mu, \sigma)$ converges P_λ a.s. to δ_{α_k}, the Dirac measure on α_k. Proceeding, we define $M_k \subset P(\Lambda)$ by $M_k = \{\mu \in P(\Lambda)$: (i) μ has finite support, (ii) $\mu(\{\alpha_k\}) > 0$, and (iii) $\mu(\Lambda_0) = 1 - \mu(\{\alpha_k\})\}$. The set $M \subset P(\Lambda)$ is defined by $M = \cup_{k=1}^{\infty} M_k$.

LEMMA 3.2. *For* $k = 1, 2, \ldots$, M_k *is a dense subset of* $P(\Lambda)$.

Proof. Select $\alpha_K \in D$, $\mu \in P(\Lambda)$ and define $\Xi = \{\gamma \in P(\Lambda)$: supp γ is finite, and $\gamma(\Lambda_0) = 1\}$. Since Λ_0 is dense in Λ (Lemma 2.1), by Theorem II.6.3 of Parthasarathy (1967), Ξ is dense in $P(\Lambda)$. So there exists $\gamma^m \Rightarrow \mu$, with $\gamma^m \in \Xi$. Define $\mu^m = m^{-1} \cdot \delta_{\alpha_K} + (1 - m^{-1}) \cdot \gamma^m$. Since $\mu^m \in M_k$ and $\mu^m \Rightarrow \mu$, the proof is complete. ■

Given the above definitions, it is intuitive that if the prior $\mu \in M_k$ and $\lambda_0 \in \Lambda_+$, then with P_{λ_0} probability one, any a priori alternative to λ_0 will eventually be deemed impossible and the posterior belief will converge to δ_{α_k}.

LEMMA 3.3. *For* $\lambda_0 \in \Lambda_+$ *and* $\mu \in M_k$, $P_{\lambda_0}(\{\sigma \in \Sigma: \Gamma_n(\mu, \sigma) \Rightarrow \delta_{\alpha_K}\}) = 1$.

Proof. Suppose supp $\mu = \{\alpha_k, \lambda_1, \lambda_2, \ldots, \lambda_J\}$ where $\lambda_j \in \Lambda_0$ for $j = 1, \ldots, J$. Let $A_j = \{z \in Z: h(\lambda_j, z) = 0\}$. Define the exceptional set $E_j = \{\sigma = (z_1, z_2, \ldots) \in \Sigma: z_k \notin A_j$ for $k = 1, 2, \ldots\}$. $P_{\lambda_0}(E_j) = 0$, and for $\sigma \notin E_j$, $\Gamma_n(\mu, \sigma)(\{\lambda_j\}) > 0$ only finitely often. Since $\{\lambda_1, \ldots, \lambda_J\}$ is finite, $\Gamma_n(\mu, \sigma)(\{\alpha_k\}) < 1$ only finitely often. ■

Now for k = 1, 2, ... , let $\{O_{km}\}_{m=1}^{\infty}$ be a decreasing sequence of open subsets of Λ with $\bar{O}_{k,m+1} \subset O_{km}$ and $\cap_{m=1}^{\infty} O_{km} = \{\alpha_k\}$. Let $g_{km}: \Lambda \to \mathbf{R}$ be a bounded Lipschitz function such that $\|g_{km}\|_{BL} \leq 1$, g_{km} equals one on $O_{k,m+1}$ and vanishs on $\sim O_{km}$. The existence of functions satisfying these conditions follows from Proposition 11.2.3 of Dudley (1989).

LEMMA 3.4. *If* $\lambda_0 \in \Lambda_+$ *and* $\mu \in M_k$, *then* $\lim_{n\to\infty} \int_\Omega [\int_\Lambda g_{km}(\lambda)\, \Gamma_n(\mu, \sigma)(d\lambda)]\, P_{\lambda_0}(d\sigma) = 1.$

Proof. Define $G_{kmn}(\sigma) = \int_\Lambda g_{km}(\lambda)\Gamma_n(\mu, \sigma)(d\lambda)$. Applying Lemma 3.3 and the definition of weak convergence, $\lim_{n\to\infty} G_{kmn}(\sigma) = 1$, P_{λ_0} a.s.. So by the Dominated Convergence Theorem, $\lim_{n\to\infty} \int_\Sigma G_{kmn}(\sigma)\, P_{\lambda_0}(d\sigma) = 1.$ ■

For arbitrary prior beliefs $\mu \in P(\Lambda)$ and 'true' parameter $\lambda \in \Lambda$, the probability law (with respect to the measureP_λ) of the posterior mapping $\Gamma_n(\mu, \cdot): \Sigma \to P(\Lambda)$ may vary with the choice of versions of conditional probability. It is easily confirmed, however, that if $\mu \in P_+(\Lambda)$ then with P_λ probability one, any two versions of conditional probability will agree. The next task is to verify that if we restrict attention to prior beliefs $\mu \in P_+(\Lambda)$, then from the perspective of statistical observer who "knows" λ, the expected value of the Bayesian's posterior expectation of g_{km} is a continuous function of the prior μ and the true parameter λ. This is a lengthy exercise with the details provided in the Appendix. Since $P_+(\Lambda)$ is a residual set, for the purposes of this paper this restricted continuity result suffices.

PROPOSITION 3.5. *The function* $\Phi_{kmn}: \Lambda \times P_+(\Lambda) \to \mathbf{R}$ *defined by*
$\Phi_{kmn}(\lambda, \mu) = \int_\Sigma \int_\Lambda g_{km}(\lambda')\, \Gamma_n(\mu, \sigma)(d\lambda')\, P_\lambda(d\sigma)$, *is continuous for all k, m and n.*

Proof. See Appendix. ■

The remaining steps needed for the proof of Theorem 3.11 mimic Freedman (1965). To make the paper self-contained, modulo notational changes (and filling in some details) we replicate Freedman's ingenious argument.

Define for k, j, m, n = 1, 2, ... , the set $R_{kmjn} \subset \Lambda \times P_+(\Lambda)$ by:

$$R_{kmjn} = \{(\lambda, \mu) \in \Lambda \times P_+(\Lambda): \int_\Sigma [\int_\Lambda g_{km}(\lambda')\, \Gamma^n(\mu, \sigma)(d\lambda')]\, P_\lambda(d\sigma) \leq 1 - j^{-1}\}.$$

And define $\mathscr{S} = \cup_{k=1}^{\infty} \cup_{m=1}^{\infty} \cup_{j=1}^{\infty} \cap_{n=1}^{\infty} R_{kmjn}$. The set of (λ, μ) pairs such that with P_λ probability one, the Bayesian's posterior belief essentially concentrates in every open Λ set infinitely often is:

$$\mathscr{R} = \{(\lambda, \mu) \in \Lambda \times P(\Lambda): \limsup_{n\to\infty} \int_\Sigma \Gamma_n(\mu, \sigma)(G)\, P_\lambda(d\sigma) = 1, \forall \text{ open } G \subset \Lambda\}.$$

To aid the reader, we outline the structure of the remainder of the proof. In Proposition 3.6 and 3.7 we establish that: (i) $\mathscr{S}$ is of first category in $\Lambda \times P_+(\Lambda)$, and (ii) that $[\Lambda \times P_+(\Lambda)]\backslash\mathscr{R} \subset \mathscr{S}$ implying that $\mathscr{R}$ is

residual in $[\Lambda \times P_+(\Lambda)]$. In conjunction with the fact that a residual subset of a residual subspace is residual (a consequence of Lemma 3.9), this establishes that $\mathcal{R}$ is residual in $\Lambda \times P(\Lambda)$.

PROPOSITION 3.6. *For all $k, m, j \geq 1$, $\cap_{n=i}^{\infty} R_{kmjn}$ is a relatively closed, nowhere dense subset of $\Lambda \times P_+(\Lambda)$. $\mathcal{F}$ is of first category in $\Lambda \times P_+(\Lambda)$.*

Proof. By Proposition 3.5, R_{kmjn} is closed in $\Lambda \times P_+(\Lambda)$ and so $\cap_{n=i}^{\infty} R_{kmjn}$ is closed in $\Lambda \times P_+(\Lambda)$. By Lemma 3.4, if $(\lambda, \mu) \in \Lambda_+ \times M_k$ then $(\lambda, \mu) \notin \cap_{n=i}^{\infty} R_{kmjn}$. By Lemma 3.2 $\Lambda_+ \times M_k$ is dense in $\Lambda \times P_+(\Lambda)$, and so $\cap_{n=i}^{\infty} R_{kmjn}$ is nowhere dense in $\Lambda \times P_+(\Lambda)$. Since $\cap_{n=i}^{\infty} R_{kmjn}$ is a closed, nowhere dense subset of $\Lambda \times P_+(\Lambda)$, $\mathcal{F}$ is a countable union of closed, nowhere dense sets, and so is of first category in $\Lambda \times P_+(\Lambda)$. ■

PROPOSITION 3.7. $[\Lambda \times P_+(\Lambda)] \setminus \mathcal{R} \subset \mathcal{F}$

Proof. Suppose $(\lambda, \mu) \in [\Lambda \times P_+(\Lambda)] \setminus \mathcal{R}$. Then there exists $\varepsilon > 0$ and an open set $G \subset \Lambda$, such that $\limsup_n \int_\Omega \Gamma_n(\mu, \sigma)(G)\ P_\lambda(d\sigma) < 1 - \varepsilon$. Choose k and m so that $\{\lambda' : g_{km}(\lambda') > 0\} \subset G$. Choose j so that $1 - j^{-1} > 1 - \varepsilon$ and choose n_0 so that for all $n \geq n_0$, $\int_\Omega \Gamma_n(\mu, \sigma)(G)\ P_\lambda(d\sigma) < 1 - j^{-1}$. Then $(\lambda, \mu) \in R_{kmjn}$ and so $(\lambda, \mu) \in \cap_{n=n_0}^{\infty} R_{kmjn}$, which by the definition of $\mathcal{F}$ implies that $(\lambda, \mu) \in \mathcal{F}$. ■

COROLLARY 3.8. *$\mathcal{R} \cap [\Lambda \times P_+(\Lambda)]$ is a residual subset of $\Lambda \times P_+(\Lambda)$ in the relative topology.*

Proof. By Propositions 3.6 and 3.7, $[\Lambda \times P_+(\Lambda)] \setminus \mathcal{R}$ is of first category in $\Lambda \times P_+(\Lambda)$, so the relative complement $R \cap [\Lambda \times P_+(\Lambda)]$ is residual in $\Lambda \times P_+(\Lambda)$. ■

LEMMA 3.9. *Suppose (Y, T) is a topological space and (Y_0, U) is a subspace where U is the relative topology.*
(i) If $A \subset Y_0$ is nowhere dense in (Y_0, U) then A is nowhere dense in (Y, T).
(ii) If Y_0 is a residual subset of Y and $B \subset Y_0$ is residual in (Y_0, U) then B is a residual subset of Y.

Proof. (i) Suppose A is not nowhere dense in Y. Then there is an open set $G \subset Y$ with $G \subset \bar{A}$. Furthermore, $G \cap Y_0 \subset \bar{A} \cap Y_0 = \mathrm{cl}_{Y_0} A$, the closure of A in (Y_0, U), where the equality follows from Theorem 1.16 of Kelley (1985). But this contradicts A being nowhere dense in Y_0.
(ii) To establish that B is residual in Y, observe that there exists a sequence of sets $\{A_i\}$ with $A_i \subset Y_0$, $Y_0 \setminus B = \cup_i A_i$, and A_i nowhere dense in (Y_0, U). So by (i), A_i is nowhere dense in Y, implying that $Y_0 \setminus B$ is of first category in Y, so $Y_0 \setminus B \cup Y \setminus Y_0$ is of first category in Y, and $B = \sim(Y_0 \setminus B \cup Y \setminus Y_0)$ is residual in Y. ■

Proof of THEOREM 3.10. By Lemma 3.1, $P_+(\Lambda)$ is a residual subset of $P(\Lambda)$. So by Theorem 15.3 of Oxtoby (1980), $\Lambda \times P_+(\Lambda)$ is a residual subset of $\Lambda \times P(\Lambda)$. By Corollary 3.8, $\mathcal{R} \cap [\Lambda \times P_+(\Lambda)]$ is a residual subset of $\Lambda \times P(\Lambda)$, and so by Lemma 3.9, $\mathcal{R} \cap [\Lambda \times P_+(\Lambda)]$ is a residual subset of $\Lambda \times P(\Lambda)$. This completes the proof that

$\mathscr{B}$ is a residual subset of $\Lambda \times P(\Lambda)$. Applying Theorem 3.13 of Dubins and Freedman (1964), S is a residual subset of $P(\Lambda)$, and by Oxtoby (1980, Theorem 15.3) $\Lambda \times S$ is a residual subset of $\Lambda \times P(\Lambda)$. So $\mathscr{B} \cap (\Lambda \times S) = \mathscr{B}_S$ is residual. ■

Proof of COROLLARY 3.11. Pick $(\lambda, \mu) \in \mathscr{B}$, $\varepsilon > 0$, and choose a set $G \subset \Lambda$ open in the $\mathscr{T}_w$ topology with $\mathscr{T}_w$-closure $\bar{G}$ such that $\lambda \notin \bar{G}$. Since $(\lambda, \mu) \in \mathscr{B}$, P_λ a.s. $\Gamma_n(\mu, \sigma)(G) > 1 - \varepsilon$ infinitely often, $\Gamma_n(\mu, \sigma)(\sim \bar{G}) < \varepsilon$ i.o.. But by the standard characterization of weak convergence (see e.g., Billingsley (1968, Theorem 2.1)) this implies that $\Gamma_n(\mu, \sigma)$ does not converge $\overset{w}{\Rightarrow}$ to δ_λ. And since $\mathscr{K}_w$ is weaker than $\mathscr{K}_1$, $\Gamma_n(\mu, \sigma)$ does not converge $\Rightarrow$ to δ_λ. ■

4. AN APPLICATION: INFINITE HORRIZON BANDIT PROBLEMS WITH DISCOUNTING

4.1. *Introduction*

In this section we model a Bayesian decision-maker who faces an infinite-horizon two-armed bandit problem and geometrically discounts future rewards. In a well-known article Rothschild (1974) applied the bandit framework to model the decision-making of a monopolist who could charge one of two prices and was uncertain of the distribution of demand associated with each price. An extended discussion of economic applications of bandit problems in provided by Kiefer (1989)

To orient the reader we first provide an informal description of the Bayesian's optimization problem. We then formally define the relevant probability spaces and reformulate the decision-problem as a dynamic programming problem. Using the Gittens Index, we provide a simple characterization the behavioral rules of the decision-maker. In Section 5, we apply Theorem 3.10 to describe the asymptotic behavior of the decision-maker.

Time periods are indexed by $t = 0, 1, 2, \ldots$. In period t the decision-maker selects an *action* or *bandit-arm* $x(t) \in X = \{x_1, x_2\}$. After choosing action $x(t) = x_k$, the realization $y(t) \in Y_k$ of a random element $Y(t)$ is observed, and a period reward $r_k(y(t))$ is received. Conditional upon the Bayesian choosing $x(t) = x_k$, the probability distribution of $Y(t)$ is an element of $\theta_k \in \Theta_k \subset P(Y_k)$. Defining $Y = Y_1 \cup Y_2$, and $r: X \times Y \to \mathbf{R}$ by $r(x_k, y) = r_k(y)$, the total reward or utility from the stream $(x(0), y(0), x(1), y(1), x(2), \ldots)$ is $\sum_{t=0}^{\infty} \alpha^t \cdot r(x(t), y(t))$, where the discount factor $\alpha \in [0, 1)$.

The decision-maker may initially be uncertain of the "true" θ_1 and θ_2. Defining $\Theta = \Theta_1 \times \Theta_2$, her initial beliefs are given by a prior probability $\mu \in P(\Theta)$. Since the decion-maker's choice of action at time t may be influenced by previously observed random outcomes, the action $x(t)$ is the realization of a random variable $X(t)$. A *policy* is a sequence of random variables $\{X(t)\}$ taking values in X, that are measurable with respect to the information (i.e., sub-σ-fields) generated by past outcomes. The objective of the decision-maker is to select a policy that maximizes $E_\mu[\sum_{t=1}^{\infty} \alpha^{t-1} \cdot r(X(t), Y(t))]$, where the symbol E_μ (informally) denotes expectation with respect to a probability measure on the underlying probability space consistent with the prior μ.

4.2. *Formal Specification of the Probability Spaces*

The set Y_k of outcomes that might arise from selecting action x_k is a non-finite, complete, separable metric space with Borel σ-field $B(Y_k)$. Define $\Omega_k = Y_k \times Y_k \times \dots$ with F_k the corresponding product σ-field. We write F_{kn} for the n-fold product σ-field $B(Y_k) \times B(Y_k) \times \dots \times B(Y_k)$ with $F_{k0} = \{\Omega_k, \varnothing\}$. The measurable space (Ω, F) is defined by $\Omega = \Omega_1 \times \Omega_2$, and $F = F_1 \times F_2$. The projection function π_k: $N \times \Omega_k \to Y_k$ is defined by $\pi_k(j, \omega_{k1}, \omega_{k2}, \dots) = \omega_{kj}$. The interpretation of Ω_k is that for $\omega_k \in \Omega_k$, $\pi_k(j, \omega_k)$ is the outcome on the j'th occasion that action x_k is selected. Slightly abusing notation, when convenient we treat $\pi_k(j, \cdot)$ as a function from Ω to Y defined by $\pi_k(j, \omega_1, \omega_2) = \pi_k(j, \omega_{kj})$.

The set of a priori possible probability laws on (Y_k, F_k) is $\Theta_k = \{\theta_k \in P(\Theta_k): \theta_k << v_k\}$, where v_k is a σ-finite measure on $(Y_k, B(Y_k))$ with non-finite support. The reward function r_k: $Y_k \to \mathbf{R}$ is a bounded, measurable function with ess inf $r_k(y_k) = \sup\{a \in \mathbf{R}: v_k(\{y_k: r_k(y_k) < a\}) = 0\} = b_k$, and ess sup $r_k(y_k) = \inf\{a \in \mathbf{R}: v_k(\{y_k: r_k(y_k) > a\}) = 0\} = c_k$. We define $\|r_k\| = \max\{|b_k|, |c_k|\}$.

With d_k the L^1 metric on Θ_k, (Θ_k, d_k) is a complete, separable metric space. For $\theta_k \in \Theta_k$, the product measure $\theta_k \times \theta_k \times \dots$ on (Ω_k, F_k) is denoted by P_{k,θ_k}. The product space (Θ, d_Θ) is defined by $\Theta = \Theta_1 \times \Theta_2$ where d_Θ is a metric that metrizes the product toplogy. The Borel σ-field is $B(\Theta) = B(\Theta_1) \times B(\Theta_2)$. For $\theta = (\theta_1, \theta_2) \in \Theta$, the product probability measure P_θ on (Ω, F) is defined by $P_\theta = P_{1,\theta_1} \times P_{2,\theta_2}$.

In order to make use of Gittens Index machinery, we require that observing an outcome from arm k provides no information regarding the probability law governing arm k′, for $k \neq k'$. More formally the prior belief must be a product probability on $(\Theta, B(\Theta))$. Define $\Pi(\Theta) = \{\mu \in P(\Theta): \mu = \mu_1 \times \mu_2$, s.t. $\mu_1 \in P(\Theta_1), \mu_2 \in P(\Theta_2)\}$. Identifying $\mu = \mu_1 \times \mu_2 \in \Pi(\Theta)$ by the vector (μ_1, μ_2) of marginals, no confusion should arise from writing (μ_1, μ_2) instead of $\mu_1 \times \mu_2$. The projection functions ρ_k: $\Pi(\Theta) \to P(\Theta_k)$ are defined by $\rho_k((\mu_1, \mu_2)) = \mu_k$.

Corresponding to a prior μ are induced probability measures P_μ and Q_μ on respectively (Ω, F) and $(\Theta \times \Omega, B(\Theta) \times F)$. For $\mu = (\mu_1, \mu_2) \in \Pi(\Theta)$, Q_μ is defined on measurable rectangles by $Q_\mu(A \times B) = \int_A P_\theta(B)\, \mu(d\theta)$ for $A \in B(\Theta)$ and $B \in F$. By the Product Measure Theorem (see Ash (1972, Theorem 2.6.2)) there is a unique extension onto $B(\Theta) \times F$. P_μ is defined by $P_\mu(B) = P_\mu(\Theta \times B)$ for every $B \in F$. From the perspective of the Bayesian decision-maker who has prior μ, the relevant probability spaces are (Ω, F, P_μ) and $(\Theta \times \Omega, B(\Theta) \times F, Q_\mu)$. But from the perspective of a classical statistician, the probability space on which all random variables are defined is (Ω, F, P_θ) where $\theta = (\theta_1, \theta_2) \in \Theta$ is the "true" parameter.

4.3. *Histories and Policies*

We now provide a precise statement of the optimization problem from the perspective of the decision-maker. We start by defining an admissible plan. It will be convenient to include in our definition, "count functions"

$C_k(t)$: $\Omega \to \{0, 1, 2, \dots\}$, for k = 1, 2 and t = 0, 1, 2, … . The realization $C_k(t)(\omega)$ will be interpretable as the number of occasions that arm k has been chosen through time t.

The indicator function I_k: $X \to \{0, 1\}$ is defined by $I_k(x) = 1$ iff $x = x_k$. An *admissible plan* is a tuple $[\{X(t)\}, \{Y(t)\}, \{C_1(t)\}, \{C_2(t)\}, \{H_t\}]$ where for t = 0, 1, 2, … , $X(t)$: $\Omega \to X$, $Y(t)$: $\Omega \to Y$, $C_k(t)$: $\Omega \to \{0, 1, \dots\}$, and $\{H_t\}$ is a sequence of sub-σ-fields of F, such that:

i) $H_0 = \{\emptyset, \Omega\}$

ii) $X(t)$: $\Omega \to X$ is H_t measurable,

iii) for $t \geq 1$, $H_t = H_{t-1} \vee \sigma(X(t-1), Y(t-1))$,

iv) $C_k(0) = I_k(X(0))$,

v) $C_k(t) = C_k(t-1) + I_k(X(t))$ for $t \geq 1$, and

vi) if $X(t)(\omega) = x_k$ and $C_k(\omega) = n$, then $Y(t)((\omega_1, \omega_2)) = \pi_k(n, \omega_k)$.

For $\alpha \in (0, 1)$, an admissible plan $[\{X(t)\}, \{Y(t)\}, \{C_1(t)\}, \{C_2(t)\}, \{H_t\}]$ is *(μ, α)–optimal* if for any other admissible plan $[\{X(t)'\}, \{Y(t)'\}, \{C_1(t)'\}, \{C_2(t)'\}, \{H_t'\}]$:

$$\int_\Omega [\sum_{t=1}^{\infty} \alpha^{t-1} \cdot r(X(t)(\omega), Y(t)(\omega))]\, P_\mu(d\omega) \geq \int_\Omega [\sum_{t=1}^{\infty} \alpha^{t-1} \cdot r(X(t)'(\omega), Y(t)'(\omega))]\, P_\mu(d\omega).^2$$

4.4. *Reformulation as a Bayesian Dynamic Programming Problem*

In this section we recast the decision-maker's optimization problem as a Bayesian dynamic programming problem with the state at time t being the Bayesian's posterior probability $\mu(t) = (\mu_1(t), \mu_2(t)) \in \Pi(\Theta)$. As in Blackwell (1965), a *dynamic programming problem* is defined as a quintuple $[X, u, \Pi(\Theta), \varphi, \alpha]$, where the *action space* X, the *discount factor* $\alpha \in [0, 1)$ and the *state space* $\Pi(\Theta)$ are as previously defined. The expected reward function for arm k is u_k: $P(\Theta_k) \to \mathbf{R}$ given by $u_k(\mu_k) = \int_{\Theta_k} \int_{Y_k} r_k(y_k)\, \theta_k(dy_k)\, \mu_k(d\theta_k)$. The *expected reward function* u: $\Pi(\Theta) \times X \to \mathbf{R}$, is defined by $u((\mu_1, \mu_2), x_k) = u_k(\mu_k)$.

To define the *transition probability* φ: $\Pi(\Theta) \times X \to P(\Pi(\Theta))$, we need to develop some notation for Bayesian updating. Analogous to the definition of Γ in Section III, the arm k updating maps Γ_k: $P(\Theta_k) \times Y_k \to P(\Theta_k)$, k = 1, 2, are chosen so that: (i) Γ_k is jointly measurable, and (ii) for each $\mu_k \in P(\Theta_k)$, $\Gamma_k(\mu_k, \cdot)$ is a regular version of conditional probability. The updating maps Γ_{kn}: $P(\Theta_k) \times \Omega_k \to P(\Theta_k)$ are recursively defined by: (i) $\Gamma_{k1}(\mu_k, \omega_k) = \Gamma_k(\mu_k, \pi_k(1, \omega_k))$ and, (ii) for n > 1, $\Gamma_{kn}(\mu_k, \omega_k) = \Gamma_k(\Gamma_{k,n-1}(\mu_k, \omega_k), \pi_k(n, \omega_k))$.

Given that action x_k is chosen with prior marginal belief μ_k, the Bayesian's probability distribution over next period marginal posterior beliefs is $\Psi_k(\mu_k) \in P^2(\Theta)$, where the map Ψ_k: $P(\Theta_k) \to P^2(\Theta_k)$ is constructed as follows. For $\mu_k \in P(\Theta_k)$ and $B \in B(P(\Theta_k))$, define $\upsilon(\mu_k, B) = \Gamma_k^{-1}(\mu_k, \cdot)(B)$. Since $\upsilon(\mu_k, B) \in B(Y_k)$, the map

$\theta_k \to \theta_k(\upsilon(\mu_k, B))$ is measurable and $\Psi_k(\mu_k)(B) = \int_{\Theta_k} \theta_k(\upsilon(\mu_k, B))\, \mu_k(d\theta_k)$. We can now define the transition probability $\varphi: \Pi(\Theta) \times X \to P(\Pi(\Theta))$ by:

$$\varphi((\mu_1, \mu_2), x_1)(A_1 \times A_2) = \Psi_1(\mu_1)(A_1) \cdot I_{A_2}(\mu_2),$$

and

$$\varphi((\mu_1, \mu_2), x_2)(A_1 \times A_2) = \Psi_2(\mu_2)(A_2) \cdot I_{A_1}(\mu_1),$$

for $A_1 \in B(P(\Theta_1))$, $A_2 \in B(P(\Theta_2))$ and $I_{A_k}(\cdot)$ the indicator function.

A *policy* is a sequence $\{X(t)\}$ such that there exists a (unique) admissible plan $[\{X(t)\}, \{Y(t)\}, \{C_1(t)\}, \{C_2(t)\}, \{H_t\}]$. The set of policies is denoted by Ξ. Given $\{X(t)\} \in \Xi$ and prior belief $\mu(0) = (\mu_1(0), \mu_2(0)) \in \Pi(\Theta)$, there is an induced sequence $\{\mu(t)\} = \{\mu_1(t), \mu_2(t)\}$ of *state* variables (posterior beliefs), where $\mu_k(t): \Omega \to P(\Theta_k)$. For $t \geq 1$, $\mu_k(t)$ is defined by: (i) if $X(t\text{-}1)(\omega) = x_k$, then $\mu_k(t)(\omega) = \Gamma_k(\mu_k(t\text{-}1), \pi_k(C_k(t), \omega))$, and (ii) if $X(t\text{-}1)(\omega) \neq x_k$, then $\mu_k(t)(\omega) = \mu_k(t\text{-}1)(\omega)$.

For $\alpha \in (0, 1)$ and $\mu(0) \in \Pi(\Theta)$, a policy $\{X(t)\}$ with corresponding posterior sequence $\{(\mu(t)\}$ is $(\mu(0), \alpha)$ *DP-optimal* if for any policy $\{X'(t)\}$ with corresponding posterior sequence $\{\mu'(t)\}$:

$$\int_\Omega \sum_{t=0}^{\infty} \alpha^t \cdot u(\mu(t)(\omega), X(t)(\omega))\, P_{\mu(0)}(d\omega) \geq \int_\Omega \sum_{t=0}^{\infty} \alpha^t \cdot u(\mu'(t)(\omega), X'(t)(\omega))\, P_{\mu(0)}(d\omega).$$

A seemingly obvious, but non-trivial fact is that for $\alpha \in (0, 1)$ an admissible plan $[\{X(t)\}, \{Y(t)\}, \{C_1(t)\}, \{C_2(t)\}, \{H_t\}]$ is (μ, α) optimal (as defined in Section 4.3) iff $\{X(t)\}$ is (μ, α) DP-optimal. Formally, this result follows from Theorem 7.3 of Reider (1975).

For $\alpha = 0$, the above definition of optimality would not useful since it would impose no restrictions on behavior after period 0. Associating $\alpha = 0$ with repeatedly myopic behavior, we define a policy $\{X(t)\}$ with corresponding posterior sequence $(\mu(t)\}$ to be $(\mu(0), 0)$ DP-optimal if for all t, $u(\mu(t), X(t)) \geq u(\mu(t), x_k)$, $P_{\mu(0)}$ a.s..

A policy $\{X(t)\}$ with associated posterior maps $\{\mu_1(t)\}$, $\{\mu_2(t)\}$ is *stationary* if there exists a *policy function* $\xi: \Pi(\Theta) \to X$ such that ξ is measurable and for all t, $X(t) = \xi(\mu_1(t), \mu_2(t))$. If $\{X(t)\}$ is stationary and DP-optimal, then the policy function ξ is an *optimal policy function*. Applying standard results in dynamic programming (see e.g., Blackwell (1965) or Maitra (1968)) it is routine to verify that an optimal, stationary policy exists.

We now restate some standard dynamic programming results in the context of our model. For discount factor α, the *value function* $V: P(\Theta) \to R$ is defined by

$$V(\mu) = \sup_{\{X(t)\} \in \Xi} \left\{ \int_\Omega \sum_{t=0}^{\infty} \alpha^t \cdot u(\mu(t)(\omega), X(t)(\omega))\, P_\mu(d\omega) \right\}$$

where $\{\mu(t)\}$ denotes the posterior sequence corresponding to $\{X(t)\}$. V satisfies the Bellman or optimality equation (see e.g. Blackwell (1965, Theorem 6 (e))):

$$V(\mu) = \mathrm{Max}_{x \in X} \left\{ u(\mu, x) + \alpha \cdot \int_{\Pi(\Theta)} V(\mu')\, \varphi(\mu, x)(d\mu') \right\}.$$

Additionally, ξ is an optimal policy function iff it solves the optimality equation (Blackwell [1965, Theorem 6 (f)]). That is for all $\mu \in \Pi(\Theta)$:

$$V(\mu) = u(\mu, \xi(\mu)) + \alpha \cdot \int_{\Pi(\Theta)} V(\mu') \, \varphi(\mu, \xi(\mu))(d\mu').$$

4.5. *Gittens Index Characterization of Optimal Policy*

There is a simple characterization of optimal policies, based upon initial results of Gittens and Jones (1974), and subsequent refinements by Berry and Fristedt (1985), Ross (1983) and Whittle (1982). Gittins and Jones proved the existence of functions M_1: $P(\Theta_1) \to R$ and M_2: $P(\Theta_2) \to R$ with the property that it is optimal to choose arm 1 in period t iff $M_1(\mu_1(t)) \geq M_2(\mu_2(t))$. The function M_k is commonly referred to as the *Gittins Index* for arm k.

To motivate the definition of M_k consider a one-armed bandit problem where in the initial stage the decision-maker has the option of playing arm 1 or stopping and collecting a terminal reward of m. In subsequent stages, assuming arm 1 has been played in all previous stages, the options remain selecting arm 1 or stopping and receiving a final payment m. The value of the Gittins Index for belief μ_k is the terminal reward m such that the decision-maker is indifferent between between continuing and stopping.

More precisely, denote the set of stopping times on $(\Omega_k, F_k\}$ as $\mathcal{T}_k = \{\tau: \Omega_k \to N_0, \tau^{-1}(\{n\}) \in F_{kn}\}$. The expected total reward of the stopping time policy τ with terminal payoff m and belief $\mu_k \in P(\Theta_k)$ is given by the function T_k: $\mathcal{T}_k \times P(\Theta_k) \times R \to R$, defined by

$$T_k(\tau, \mu_k, m) = \int_{\Omega_k} [[\sum_{t=0}^{\tau(\omega_k)-1} \alpha^t \cdot r_k(\omega_{k,t+1})] + \alpha^{\tau(\omega_k)} \cdot m] \, P_{k,\mu_k}(d\omega_k).$$

The value of the optimal policy is $V_k(\mu_k, m) = \mathrm{Sup}_{\tau \in \mathcal{T}_k}\{T_k(\tau, \mu_k, m)\}$. Recalling that $c_k = \text{ess sup}\{r_k(y_k)\}$, it is routine to verify that for all $\mu_k \in P(\Theta_k)$, $V_k(\mu_k, m) = m$ for all $m \geq c_k \cdot (1 - \alpha)^{-1}$. The Gittins Index is defined by $M_k(\mu_k) = \inf\{m: V_k(\mu_k, m) = m\}$. A characterization of the optimal policy in terms of the Gittins Index is given by the following proposition.

PROPOSITION 4.1. *A policy* $\{X(t)\}$ *with posterior maps* $\{\mu_1(t)\}$, $\{\mu_2(t)\}$ is $(\mu_1(0), \mu_2(0))$*-optimal* iff $X(t)(\omega) = x_k$ *whenever* $M_k(\mu_k(t)(\omega)) > M_j(\mu_j(t)(\omega))$.

Proof. Whittle (1982, Theorem 14.4.1). ■

Throughout the remainder of the paper we assume that the Bayesian controller follows an optimal stationary policy $\{X(t)\}$ with posterior maps $\{\mu_1(t)\}$, $\{\mu_2(t)\}$, and policy function ξ: $\Pi(\Theta) \to X$ defined by $\xi((\mu_1(t), \mu_2(t))) = x_1$ iff $M_1(\mu_1(t)) \geq M_2(\mu_2(t))$. For a more detailed development of the Gittens Index and the optimality of the Index policy, the texts of Ross (1983), Whittle (1982), and Berry and Fristedt (1985) are recommended.

5. GENERIC LIMIT THEOREMS

5.1. Some Continuity Results

Preparatory to proving the genericity theorems, some preliminary technical results are developed in this subsection. The principal result is the establishment of the continuity of the functions M_1 and M_2. The first step is to develop characterizations of V_1 and V_2 by making use of the fact that these value functions are solutions to the optimality equation.

LEMMA 5.1. *The function* $V_k: P(\Theta_k) \times \mathbf{R} \to \mathbf{R}$ *is continuous, for* $k = 1, 2$.

Proof. Without loss of generality we consider the case of $k = 1$. We first note that an easy consequence of Lemma A.5. is that Ψ_1 is continuous. Let $C(P(\Theta_1))$ denote the space of bounded, continuous real-valued functions on $P(\Theta_1)$ endowed with the sup norm. Since $u(\mu_1, \mu_2, x_1)$ is independent of μ_2, we may define $u_1(\mu_1) = u(\mu_1, \mu_2, x_1)$. For each $m \in \mathbf{R}$, define the mapping $\vartheta_m: C(P(\Theta_1)) \to C(P(\Theta_1))$ by $\vartheta_m(\zeta)(\mu_1) = \max\{m, u_1(\mu_1) + \alpha \cdot \int \zeta(\gamma_1)\, \Psi_1(\mu_1)(d\gamma_1)\}$. (The continuity of $\vartheta_m(\zeta(\cdot))$ follows from the continuity of u_1 and Ψ_1.) By standard arguments one can verify that ϑ_m is a contraction mapping with modulus α. Therefore the fixed-point of ϑ_m is a continuous function of m. Since $V_1(\cdot, m)$ is the fixed-point of Ξ_m, the map $(\mu_1, m) \to V_k(\mu_1, m)$ is continuous. ■

LEMMA 5.2. *For* $k = 1, 2$, *and* $\mu_k \in P(\Theta_k)$, *the map* $m \to [V_k(\mu_k, m) - m]$ *is decreasing.*

Proof. Replacing sums with integrals, the proof of Ross (1983, Lemma VII.2.1) remains valid. ■

Recall that $\Psi_k(\mu_k)$ is the Bayesian's probability distribution over next period beliefs on Θ_k, given that arm k is chosen and μ_k is the current belief. Define $W_k: P(\Theta_k) \times \mathbf{R} \to \mathbf{R}$ by $W_k(\mu_k, m) = u_k(\mu_k) + \alpha \cdot \int_{P(\Theta_k)} V_k(\mu_k', m)\, \Psi_k(\mu_k)(\mu_k')$. $W_k(\mu_k, m)$ is the expected reward for the one-armed bandit problem with the option of stopping and receiving a payoff of m after the initial period.

LEMMA 5.3. (i) W_1 *and* W_2 *are continuous.* (ii) $M_k(\mu_k) < m$ *iff* $W_k(\mu_k, m) < m$.

(i) This follows from the continuity of Ψ_k and Lemma 5.1.

(ii) Suppose $W_k(\mu, \bar{m}) < \bar{m}$. By definition of W_k, $\bar{m} > u_k(\mu) + \alpha \cdot \int V_k(\mu', \bar{m})\, \Psi_k(\mu)(d\mu')$. Since V_k is continuous (Lemma 5.1) and increasing in m, ($\bar{m} - \varepsilon > u_k(\mu) + \alpha \cdot \int V_k(\mu', \bar{m} - \varepsilon)\, \varphi_k(\mu)(d\mu')$ for ε sufficiently small. But this last inequality is, by definition, equivalent to $M_k(\mu) < \bar{m} - \varepsilon < \bar{m}$.

For the converse, suppose $M_k(\bar{\mu}) = \bar{m} < m$. Since $\bar{m}$ is a terminal payoff for which the decision-maker is indifferent between continuing play and quitting, $W_k(\bar{\mu}, \bar{m}) = V_k(\bar{\mu}, \bar{m}) = \bar{m}$. From Lemma 5.2, for all $\mu \in P(\Theta_k)$,

$V_k(\mu, m) - V_k(\mu, \bar{m}) \le (m - \bar{m})$, implying that $\int[V_k(\mu', m) - V_k(\mu', \bar{m})]\, \varphi_k(\bar{\mu})(d\mu') \le (m - \bar{m})$. Finally, we have that:

$$W_k(\bar{\mu}, m) = u_k(\bar{\mu}) + \alpha\cdot\int V_k(\mu', m)\, \Psi_k(\bar{\mu})(d\mu')$$

$$= u_k(\bar{\mu}) + \alpha\cdot\{\int V_k(\mu', m)\, \Psi_k(\bar{\mu})(d\mu') + \int V_k(\mu', \bar{m})\, \Psi_k(\bar{\mu})(d\mu') - \int V_k(\mu', \bar{m})\, \Psi_k(\bar{\mu})(d\mu')\}$$

$$= \{u_k(\bar{\mu}) + \alpha\cdot\int V_k(\mu', \bar{m})\, \Psi_k(\bar{\mu})(d\mu')\} + \alpha\cdot\{\int V_k(\mu', m)\, \Psi_k(\bar{\mu})(d\mu') - \int V_k(\mu', \bar{m})\, \Psi_k(\bar{\mu})(d\mu')\}$$

$$= \bar{m} + \alpha\cdot\{\int V_k(\mu', m)\, \Psi_k(\bar{\mu})(d\mu') - \int V_k(\mu', \bar{m})\, \Psi_k(\bar{\mu})(d\mu')\}$$

$$\le \bar{m} + \alpha\cdot(m - \bar{m})$$

$$< m. \blacksquare$$

PROPOSITION 5.4. M_1 *and* M_2 *are continuous.*

Proof. First we establish that M_k is lower-semicontinuous by verifying that for all $c \in \mathbf{R}$ the set $\{\mu \in P(\Theta_k)\colon M_k(\mu) > c\}$ is open. Suppose that $M_k(\bar{\mu}) > c$. By definition of M_k, $V_k(\bar{\mu}, c) - c > 0$. From the continuity of V_k (Lemma 5.1), $\exists$ an open neighborhood N of $\bar{\mu}$ such that for all $\mu \in N$, $V_k(\bar{\mu}, c) - c > 0$. Therefore, for all $\mu \in N$, $V_k(\mu, c) - c > 0$ and so $M_k(\mu) > c$.

Upper-semicontinuity is confirmed by demonstrating that the set $\{\mu \in P(\Theta_k)\colon M_k(\mu) < c\}$ is open for all $c \in \mathbf{R}$. Suppose that $M_k(\bar{\mu}) = \bar{m} < c$. By Lemma 5.3, $W_k(\bar{\mu}, c) < c$ and $\exists$ an open neighborhood J of $\bar{\mu}$ such that for $\mu \in J$, $W_k(\mu, c) < c$. But this implies that $M_k(\mu) < c$. $\blacksquare$

5.2. *Generic Outcomes when* $b_1 \neq b_2$

In this subsection we analyze the limit behavior of the decision-maker for the case where $b_1 \neq b_2$ (recall that $b_k = \text{ess inf}\{r_k(y_k)\colon y_k \in Y_k\}$). Without loss of generality we assume that $b_1 < b_2$. To motivate the next result, choose a set $Y_0 \subset Y_1$ with $v_1(Y_0) > 0$ and $\sup\{r_1(y_1)\colon y_1 \in Y_0\} < b_2$. So if at time t, the decision-maker's conditional probability of an outcome in Y_0 occurring if arm 1 is selected is sufficiently large, arm 2 will be selected regardless of $\mu_2(t)$; and so arm 2 will be selected at all times $t' \ge t$. From Theorem 3.10 we can conclude that there is a residual set $\mathscr{R}_1 \subset \Theta_1 \times P(\Theta_1)$ with the following property. If $(\theta_1, \mu_1) \in \mathscr{R}_1$, μ_1 is the prior belief on $(\Theta_1, B(\Theta_1))$, and arm 1 is played sufficiently often, the decision-maker will eventually, P_{θ_1} a.s., become sufficiently pessimistic that arm 1 will never be tested again. Consequently, arm 1 will be played only finitely often.

To formalize the above remarks, we begin by defining Y_0 as above. Choose $\mu_{10} \in P(\Theta_1)$ such that $\mu_{10}(Y_0) = 1$. For $\varepsilon > 0$ define an open neighborhood $G_\varepsilon \subset P(\Theta_1)$ of μ_{10}, by $G_\varepsilon = \{\mu_1 \in P(\Theta_1)\colon \beta(\mu_1, \mu_{10}) < \varepsilon\}$.

LEMMA 5.5. *There exists* $c > 0$, *such that for all* $\varepsilon < c$, $\sup_{\mu_1 \in G_\varepsilon}\{M_1(\mu_1)\} < \inf_{\mu_2 \in P(\Theta_2)}\{M_2(\mu_2)\}$.

Proof. First observe that: (i) $\inf\{M_2(\mu_2): \mu_2 \in P(\Theta_2)\} = \frac{b_2}{1-\alpha}$, and (ii) $M_1(\mu_{10}) < \frac{b_2}{1-\alpha}$. Since M_1 is continuous, $J = \{\mu_1 \in P(\Theta_1): M_1(\mu_1) < \frac{b_2}{1-\alpha}\}$ is open. So $\exists$ c such that for $\varepsilon < c$, $G_\varepsilon \subset J$. And so for $\mu_1 \in G_\varepsilon \subset J$, and any $\mu_2 \in P(\Theta_2)$, $M_1(\mu_1) < \frac{b_2}{1-\alpha} \leq M_2(\mu_2)$. ■

PROPOSITION 5.6. *Suppose* $b_1 < b_2$. *Then there exists a residual subset* $\mathscr{R}_1 \subset \Theta_1 \times P(\Theta_1)$ *such that for* $(\theta_1, \mu_1(0)) \in \mathscr{R}_1$ and all $(\theta_2, \mu_2(0)) \in \Theta_2 \times P(\Theta_2)$, P_θ *a.s.* $X(t)(\omega) = x_1$ *only finitely often.*

Proof. Invoking Lemma 5.5, choose an open set G_ε such that $M_1(\mu_1) < M_2(\mu_2)$ for all $(\mu_1, \mu_2) \in G_\varepsilon \times P(\Theta_2)$. Define the set $\Omega_{1\varepsilon} \subset \Omega_1$ by $\Omega_{1\varepsilon} = \{\omega_1 \in \Omega_1: \Gamma_{1n}(\mu_1(0), \omega_1) \notin G_\varepsilon$, for all $n = 1, 2, \ldots\}$. Define $\mathscr{R}_1 = \{(\theta_1, \mu_1(0)) \subset \Theta_1 \times P(\Theta_1): P_{\theta_1}(\Omega_{1\varepsilon}) = 0\}$. By Theorem 3.10, $\mathscr{R}_1$ is residual. ■

5.3. *Generic Outcomes when* $b_1 = b_2$.

When $b_1 = b_2$, the "typical" outcome will be that the decision-maker will start out playing one arm, but will eventually become sufficiently pessimistic regarding the first choice and switch to the other arm. Eventually, though she will become sufficiently regarding the non-initial arm and switch back to the original arm. This switching back and forth will continue forever.

PROPOSITION 5.7. *Suppose* $b_1 = b_2 = b$ and $v_k(\{y_k: r_k(y_k) \neq b\}) > 0$. *Then there is a residual set* $R \subset \Theta \times \Pi(\Theta)$ *such that* P_θ a.s., $X(t) = x_1$ *infinitely often and* $X(t) = x_2$ *infinitely often.*

Proof. Let $R_k = \{(\theta_k, \mu_k(0)) \in \Theta_k \times P_k(\Theta): P_{\theta_k}$ a.s, $\limsup_n \Gamma_{kn}(\mu_k(0), \omega_k)(G) = 1$, for all open $G \subset \Theta_k\}$. By Theorem 3.10, R_k is residual, and so $R = R_1 \times R_2$ is residual. Now choose $(\theta_1, \mu_1(0), \theta_2, \mu_2(0)) \in R$, and define $E_1 = \{\omega: \sup_t C_1(t)(\omega)$ is finite$\}$. E_1 is the ω-set for which arm 1 is played only finitely, given the optimal policy starting from beliefs $(\mu_1(0), \mu_2(0))$. Because of the symmetry of the specification, it suffices to prove that $P_\theta(E_1) = 0$. For $\omega \in E_1$, define the terminal value of the Gittens index for arm 1 as $m_{1\infty}(\omega) = M_1(\Gamma_{1,C_1(t)(\omega)}(\mu_1(0), \omega_1))$. Since $(\theta_1, \mu_1(0)) \in R_1$, $P_\theta(E_1 \cap \{\omega: m_{1\infty}(\omega) = \frac{b}{1-\alpha}\}) = 0$. Now let $E_2 = \{(\omega_1, \omega_2): \limsup_n \Gamma_{2n}(\mu_2(0), \omega_2)(G) = 1$, for all open $G \subset \Theta_2\}$, and by Theorem 3.10, $P_\theta(E_2) = 1$. By the continuity of M_2, $E_1 \subset \{E_1 \cap \{\omega: m_{1\infty}(\omega) = \frac{b}{1-\alpha}\}\} \cup \sim E_2$, and so $P_\theta(E_1) = 0$. ■

6. CONCLUDING REMARKS

In Section 3, we demonstrated the asymptotic sensitivity of posterior beliefs with respect to a prior priobability on a parameter space which is not finite dimensional. These results have significance for Bayesian statistical decision theory. Since posterior beliefs are not robust to small perturbabtions of the prior, the optimal action correspondence is similarly non-robust. If two Bayesians have identical objective or loss functions, nearby prior beliefs *and observe the same sequence of outcomes*, without additional restrictions, it would not be pathological for them to each choose

actions that the other evaluated as markedly inferior to their own choice. The bandit example, developed in Sections 4 and 5 of this paper, illustrates this phenomena in a context where beliefs do, in fact, affect outcomes.

For stochastic processes that might be generated by more complex feedback between beliefs and actions, such as rational expectations models with learning, no formal demonstration is provided of generic nonconvergence of beliefs and outcomes. Nevertheless, it would be surprising (at least to this author) if the additional complexity of such models, somehow restored the aymptotic regularity attained in analyses conducted from the perspective of the Bayesian learner.

Economic theorists who wish to model learning by Bayesian economic agents are confronted with three options. One alternative is to impose no restrictions upon agent beliefs and analyze the resulting stochastic process from the perspective of a classical statistician who knows the true parameter. With sufficiently specific assumptions on other parameters of the model (such as preferences and technology) it may be possible show that agents with consistent priors will eventually be financially dominant[3], and presumably then the limiting price process would be indistinguishable from the price process emanating from a model in which all agents had consistent priors. Even so, there would remain the question as to whether any such asymptotic properties would be robust under arbitrarily small perturbations of parameter values and initial beliefs.

A second alternative would be to follow the currently predominant practice of adopting the probabilistic vantage point of the Bayesian agent(s), imposing no restrictions on agent beliefs. A drawback with this approach is that it provides no basis for drawing distributional inferences for any particular parameter value of agent prior probability zero. The resulting limit theorems can be interpreted as predictions by the economic theorist only if the theorist's beliefs are absolutely continuous with respect to the agent's beliefs. So unless the reader also has beliefs absolutely continuous with respect to the agent's, there are no grounds for accepting the agent's asymptotic predictions.[4]

The final option, one that I endorse, is to narrow the set of candidate prior beliefs, a strategy that is adopted by Bikhchandani and Sharma (1990). To motivate this strategy, consider the bandit problem studied in Sections 4 and 5, and for simplicity suppose that Y_1 and Y_2 are each countable. If I was the decision-maker, I might find it difficult to exactly specify my prior, but I would reject any prior belief for which there was a residual θ-set A, such that for all $\theta \in A$, with P_θ probability one my beliefs would not converge. In particular, I would require that for any arm played infinitely often, the Prohorov (or bounded Lipschitz) distance between my posterior beliefs and the sample distribution converge to zero. And while acknowledging that introspective reasoning has its limitations, I believe that few individuals would behave as predicted by Propositions 5.6 and 5.7. More generally, in environments where consistent estimators are available, the modeller should assume that priors are chosen from the family of probability measures that are consistent for all $\theta \in \Theta$.

This is philosophically similar to the "what if" method advocated by Diaconis and Freedman (1986) for Bayesian statisticians. Diaconis and Freedman suggest that "... after specifying a prior distribution, generate imaginary data sequences, compute the posterior, and consider whether the posterior would be an adequate representation of the updated prior." Adapting this reccommendation to the context of economic modelling, it would be natural to require that agents have prior beliefs with full support and that the sequence of agent posterior beliefs

converges almost surely with respect to the measure P_θ for all $\theta \in \Theta$. The feasibility of such a strategy requires the existence of consistent priors. Unfortunately, endogenous learning will typically imply a non-stationary stochastic process. And there are few results currently available on the P_θ consistency (or convergence properties) of Bayes estimates for such processes.[5] Further research on sufficient conditions on priors for P_θ consistency is needed.

Appendix

For simplicity we explicitly prove Proposition 3.5 for n = 1. The argument for the n-fold product space is essentially identical and is briefly sketched.

We start by defining for $A \in B(Z)$, the function $f_A\colon \Lambda \to \mathbf{R}$ by $f_A(\lambda) = \int_A h(\lambda, z)\, v(dz)$.

LEMMA A.1. f_A *is a bounded Lipschitz function, and* $\|f_A\|_{BL} \leq 3$.

Proof. The Lipschitz norm is defined by $\|f_A\|_L = \sup_{\lambda \neq \lambda'}\{|f_A(\lambda) - f_A(\lambda')|/(d_1(\lambda, \lambda'))\}$. Since

$$|f_A(\lambda) - f_A(\lambda')| \leq \int_A |h(\lambda, z) - h(\lambda', z)|\, v(dz) \leq 2 \cdot d_1(\lambda, \lambda') \leq 2,$$

the Lipschitz norm exists and $\|f_A\|_L \leq 2$. Since $\|f_A\|_\infty = \sup_\lambda |f_A(\lambda)| = 1$, the bounded Lipschitz norm $\|f_A\|_{BL} = \|f_A\|_L + \|f_A\|_\infty$ exists and $\|f_A\|_{BL} \leq 3$. ■

LEMMA A.2. For $\mu_i, \mu \in P(\Lambda)$ if $\mu_i \Rightarrow \mu$ then $\{\sup_A[|\int_\Lambda f_A\, d(\mu_i - \mu)|: A \in B(Z)\} \to 0$.

Proof. Recall that the Dudley metric β is defined by $\beta(\mu, \gamma) = \sup\{|\int f\, d(\mu - \gamma)|: \|f\|_{BL} \leq 1\}$ and that $\mu_i \Rightarrow \mu$ iff $\beta(\mu_i, \mu) \to 0$. So by Lemma A.1, $\mu_n \Rightarrow \mu$ implies $\{\sup_A[|\int_\Lambda f_A\, d(\mu_n - \mu)|: A \in B(Z)\} \to 0$.■

For $\mu \in P(\Lambda)$ define the probability P_μ on $(Z \times \Lambda, B(Z) \times P(\Lambda))$ by $P_\mu(A \times B) = \int_B f_A(\lambda)\, \mu(d\lambda)$. Also define the marginal probability Q_μ on $(Z, B(Z))$ by $Q_\mu(A) = P_\mu(A \times \Lambda)$. Now fix a sequence $\mu_i \Rightarrow \mu \in P(\Lambda)$ and define $L_i(z) = \int_\Lambda h(\lambda, z)\, \mu_i(d\lambda)$ and $L(z) = \int_\Lambda h(\lambda, z)\, \mu(d\lambda)$. $L_i(z)$, $L(z)$ are the Radon-Nikodym derivatives of Q_{μ_j} and Q_μ with respect to v.

LEMMA A.3. $L_i(\cdot) \to L(\cdot)$ *in v measure.*

Proof. $\int_A L_i(z)\, v(dz) = \int_\Lambda \int_\Lambda h(\lambda, z)\, \mu_i(d\lambda)\, v(dz) = \int_\Lambda \int_A h(\lambda, z)\, v(dz)\, \mu_i(d\lambda)$

$$= \int_\Lambda f_A(\lambda)\, \mu_i(d\lambda) \to \int_\Lambda f_A(\lambda)\, \mu(d\lambda) = \int_A L(z)\, v(dz).$$

Furthermore, by Lemma A.2, the convergence is uniform in A. The conclusion follows from Remark 3.1 of [Strasser, 1985 #57]. ■

For $B \in B(\Lambda)$, define $L_i(B, x) = \int_B h(\lambda, z)\, \mu_i(d\lambda)$ and $L(B, z) = \int_B h(\lambda, z)\, \mu(d\lambda)$

LEMMA A.4. If $B \in B(\Lambda)$ and $\mu(\partial B) = 0$, then $L_i(B, z) \to L(B, z)$ in v measure.

Proof. First consider the case where $\mu(B) > 0$. It suffices to verify that $|\int_A L_i(B, z)\, v(dz) - \int_A L(B, z)\, v(dz)| \to 0$ uniformly in A. Let $B(B)$ denote the Borel sets restricted to B. Define $\mu_i(B) = a_i$; and define γ_i on $(B, B(B))$ by $\gamma_i(C)$

$= a_i^{-1}\cdot\mu_i(C)$, for $C \in \mathcal{B}(B)$. Similarly, define $a = \mu(B)$ and define γ on $(B, \mathcal{B}(B))$ by $\gamma(C) = a^{-1}\cdot\mu(C)$. Observe that $\int_A L_i(B, z)\, v(dz) = \int_B f_A(\lambda)\, \mu_i(d\lambda)$ and that $\int_A L(B, z)\, v(dz) = \int_B f_A(\lambda)\, \mu(d\lambda)$. So

$$\left| \int_A L_i(B, z)\, v(dz) - \int_A L(B, z)\, v(dz) \right|$$

$$= \left| \int_B f_A(\lambda)\cdot a_i\, \gamma_i(d\lambda) - \int_B f_A(\lambda)\, \gamma(d\lambda) \right|$$

$$\le \left| \int_B f_A(\lambda)\cdot \alpha_i\, \gamma_i(d\lambda) - \int_B f_A(\lambda)\, \gamma_i(d\lambda) \right| + \left| \int_B f_A(\lambda)\, \gamma_i(d\lambda) - \int_B f_A(\lambda)\, \gamma(d\lambda) \right|.$$

Since $|f_A|$ is bounded, $\left| \int_B f_A(\lambda)\cdot \alpha_i\, \gamma_i(d\lambda) - \int_B f_A(\lambda)\, \gamma_i(d\lambda) \right| \to 0$ independently of A. Substituting γ_i and γ for μ_i and μ in Lemma A.2., $\sup_{A\in\mathcal{B}(\Lambda)}\{\left| \int_B f_A(\lambda)\cdot a_i\, \gamma_i(d\lambda) - \int_B f_A(\lambda)\, \gamma(d\lambda) \right|\} \to 0$.

If $\mu(B) = 0$ then observe that $L_i(z) = L_i(B, z) + L_i(\sim B, z)$ and $L(z) = L_i(B, z) + L_i(\sim B, z)$. Since B is a μ-continuity set, $\sim B$ is also a μ-continuity set. So by application of the above and Lemma A.3, $L_i(B, \cdot) \to L(B, \cdot)$ in v measure. ■

LEMMA A.5. Suppose $\mu \in P(\Lambda)$ and pick $\varepsilon > 0$. Let $\mathcal{A} = \{A_1, A_2, \dots\}$ be a disjoint cover of Λ, with diameter of $A_t < \delta = \frac{\varepsilon}{5}$. Define $B_{t'} = \cup_{t=1}^{t'} A_t$, and choose T so that $\mu(B_T) > 1 - \delta$. If $\gamma \in P(\Lambda)$ and $\sup_{t\le T}\{|\mu(A_t) - \gamma(A_t)|\} < \delta\cdot T^{-1}$, then $\beta(\mu, \gamma) < \varepsilon$.

Proof. Let $g \in BL(\Lambda, d_1)$ with $\|g\|_{BL} \le 1$. It suffices to prove that $\left|\int_\Lambda g\, d\mu - \int_\Lambda g\, d\gamma\right| < \varepsilon$. Define $a_t = \inf\{g(\lambda)\colon \lambda \in A_t\}$, define $b_t = \sup\{g(\lambda)\colon \lambda \in A_t\}$, and note that $b_t - a_t < \delta$. Since $\int_{A_t} g\, d\gamma \ge [\mu(A_t) - \frac{\delta}{T}]\cdot a_t$, and $\int_{A_t} g\, d\mu \le [\mu(A_t)]\cdot b_t$, it follows that $\int_{A_t} g\, d\mu - \int_{A_t} g\, d\gamma \le \delta\cdot\mu(A_t) + \frac{\delta}{T}$.

Similarly, $\int_{A_t} g\, d\gamma - \int_{A_t} g\, d\mu \le \delta\cdot\mu(A_t) + \frac{\delta}{T}$. So $\left|\int_{A_t} g\, d\gamma - \int_{A_t} g\, d\mu\right| \le \delta\cdot\mu(A_t) + \frac{\delta}{T}$, and $\left|\int_{B_T} g\, d\gamma - \int_{B_T} g\, d\mu\right| \le \delta\cdot\mu(B_T) + \delta < 2\delta$.

It remains to bound the difference of the integrals over $\sim B_T$. Since $0 < \gamma(\sim B_T) < 2\delta$, $\left|\int_{\sim B_T} g\, d\mu - \int_{\sim B_T} g\, d\gamma\right| < 2\delta + \delta$. So $\left|\int_\Lambda g\, d\mu - \int_\Lambda g\, d\gamma\right| < 5\delta = \varepsilon$. ■

PROPOSITION A.6. Suppose $\mu, \mu_i \in P_+(\Lambda)$ and $\mu_i \Rightarrow \mu$. Then $\Gamma(\mu_i, z), \Gamma(\mu, z) \in P_+(\Lambda)$ v a.s., and $\Gamma(\mu_i,\cdot) \to \Gamma(\mu,\cdot)$ in v-measure.

Proof. Verification that $\Gamma(\mu_i, z), \Gamma(\mu, z) \in P_+(\Lambda)$ is routine and so omitted. Pick $\varepsilon > 0$, and as in Lemma A.5, let $\mathcal{A} = \{A_1, A_2, \dots\}$ be a disjoint cover of Λ with diameter of $A_t < \delta = \frac{\varepsilon}{5}$ and $\mu(\partial A_t) = 0$. (The existence of such a collection follows from Theorem 11.7.3. of Dudley (1989).) Define B_t as in Lemma A.5, and choose T s.t. $\mu(B_T) > 1 - \delta$. Observe that v a.s., $\Gamma(\mu_i, z)(A_t) = \frac{L_i(A_t, z)}{L_i(z)}$, which is well-defined since $\mu_i \in P_+(\Lambda)$ and so $L_i(z) > 0$.

Applying Lemmas A.3 and A.4, we have $\frac{L_i(A_l,\cdot)}{L_i(\cdot)} \to \frac{L(A_l,\cdot)}{L(\cdot)}$ in v-measure. Since B_T is the union of a finite number of A_l sets, by Lemma A.5, we have $\limsup_i v(\{z: \beta(\Gamma(\mu_i, z), \Gamma(\mu, z)) \geq \varepsilon\} = 0$. Since ε is arbitrary, $\int_Z \beta(\Gamma(\mu_i, z), \Gamma(\mu,z))\ v(dz) \to 0$. ■

PROPOSITION A.7. Suppose $\mu, \mu_i \in P_+(\Lambda)$, $\mu_i \Rightarrow \mu$, $f: \Lambda \to \mathbf{R}$ is bounded and continuous, and $\lambda_i \to \lambda \in (\Lambda, d_1)$. Then $\int_Z \int_\Lambda [f(\lambda)\ \Gamma(\mu_i, z)(d\lambda)]\ h(\lambda_i, z)\ v(dz) \to \int_Z \int_\Lambda [f(\lambda)\ \Gamma(\mu, z)(d\lambda)]\ h(\lambda, z)\ v(dz)$.

Proof. Define $\varphi_i, \varphi: Z \to \mathbf{R}$ by $\varphi_i(z) = \int_\Lambda f(\lambda)\ \Gamma(\mu_i, z)(d\lambda)$ and $\varphi(z) = \int_\Lambda f(\lambda)\ \Gamma(\mu_i, z)(d\lambda)$. By Proposition A.6, $\varphi_i \to \varphi$ in v-measure. Since $|\varphi_i|$ and $|\varphi|$ are uniformly bounded by $\|f\|$, and $h(\lambda_i, \cdot) \to h(\lambda, \cdot)$ in $L^1(Z, B(Z), v)$, $\varphi_i \cdot h(\lambda_i, \cdot) \to \varphi \cdot h(\lambda, \cdot)$ in $L^1(Z, B(Z), v)$. So $\int_Z \varphi_i(z) \cdot h(\lambda_i, z)\ v(dz) \to \int_Z \varphi(z) \cdot h(\lambda, z)\ v(dz)$, or equivalently,

$$\int_Z \int_\Lambda [f(\lambda)\ \Gamma(\mu_i, z)(d\lambda)]\ h(\lambda_i, z)\ v(dz) \to \int_Z \int_\Lambda [f(\lambda)\ \Gamma(\mu, z)(d\lambda)]\ h(\lambda, z)\ v(dz). \ ■$$

Proof of PROPOSITION 3.5. Fix k and m. For $n = 1$, setting $f = g_{kmn}$, the result follows directly from Proposition A.7. For $n > 1$, repeat the above arguments replacing z with σ and v with v^n. ■

FOOTNOTES

[1]Some authors allow for heterogeneous beliefs across agents.

[2]The definition of an optimal plan when $\alpha = 0$ will be provided in the next subsection.

[3]I am indebted to Neil Wallace for this observation.

[4]The above remarks are also applicable to nonparametric Bayesian econometrics. If the reader of a Bayesian statistical analysis has a prior not absolutely continuous with respect to the author's prior (and the prior is not consistent for all parameter values), even with large samples there may be no merging of posterior beliefs. For a more complete discussion of these issues is provided by Diaconis and Freedman (1986).

[5]Barron (1988) has recently proven results for stationary stochastic processes. Barron's techniques are potentially extendable to processes where endogenous learning generates the non-stationarity.

REFERENCES

Ash, R. (1972). "Real Analysis and Probability." New York, Academic Press.

Barron, A. (1988). "The Exponential Convergence of Posterior Probabilities with Implications for Bayes Estimators of Density Functions." Working Paper, Department of Statistics, University of Illinois at Urbana-Champaign.

Berry, D. A. and B. Fristedt. (1985). "Bandit Problems: Sequential Allocation of Experiments." London, Chapman & Hall.

Bikhchandani, S. and S. Sharma. (1990). "Optimal Search with Learning." Working Paper, UCLA.

Billingsley, P. (1968). "Convergence of Probability Measures." New York, John Wiley.

Blackwell, D. (1965). "Discounted Dynamic Programming." *Annals of Mathematical Statistics* 36, 226-235.

Diaconis, P. and D. Freedman. (1986). "On the Consistency of Bayes Estimates." *Annals of Statistics* 14, 1-26.

Dubins, L. and D. Freedman. (1964). "Measurable Sets of Measures." *Pacific Journal of Mathematics* 14, 1211-1222.

Dudley, R. M. (1966). "Convergence of Baire Measures." *Studia Mathematica* 27, 251-268.

Dudley, R. M. (1989). "Real Analysis and Probability." Belmont, CA, Wadsworth & Brooks/Cole.

Dynkin, E. B. and A. A. Yushkevich. (1979). "Controlled Markov Processes." New York, Springer-Verlag.

Easley, D. and N. Kiefer. (1988). "Controlling a Stochastic Process with Unknown Parameters." *Econometrica* (56), 1045-1064.

Easley, D. and N. Kiefer. (1989). "Optimal Learning with Endogenous Data." *International Economic Review* 30, 963-978.

Feldman, M. (1987a). "Bayesian Learning and Convergence to Rational Expectations." *Journal of Mathematical Economics* 16, 297-313.

Feldman, M. (1987b). "An Example of Convergence to Rational Expectations with Heterogeneous Beliefs." *International Economic Review* ,

Feldman, M. and A. McLennan. (1989). "Learning in a Repeated Statistical Decision Problem with Normal Disturbances."

Freedman, D. (1963). "Asymptotic Behavior of Bayes Estimates in the Discrete Case." *Annals of Mathematical Statistics* 34, 1386-1403.

Freedman, D. (1965). "On the Asymptotic Behavior of Bayes Estimates in the Discrete Case II." *Annals of Mathematical Statistics* 36, 454-456.

Gittens, J. and D. Jones. (1974). A Dynamic Allocation Index for the Sequential Design of Experiments. "Progress in Statistics." Amsterdam, North-Holland.

Kelley, J. (1985). "General Topology." New York, Springer-Verlag.

Kiefer, N. (1989). "Optimal Collection of Information by Partially Informed Agents." *Econometric Reviews* 7, 113-148.

Kiefer, N. and Y. Nyarko. (1988). Control of a Linear Regression Process with Unknown Parameters. "Third International Symposium in Economic Theory and Econometrics." Cambridge University Press.

Kiefer, N. and Y. Nyarko. (1989). "Optimal Control of an Unknown Linear Process with Learning." *International Economic Review* 30, 571-588.

Maitra, A. (1968). "Discounted Dynamic Programming in Compact Metric Spaces." *Sankhya, Ser A* 30, 211-216.

McLennan, A. (1987). "Incomplete Learning in a Repeated Statistical Decision Problem." Working Paper, University of Minnesota.

Oxtoby, J. (1980). "Measure and Category." New York, Springer-Verlag.

Parthasarathy, K. R. (1967). "Probability Measures on Metric Spaces." New York, Academic Press.

Reider, U. (1975). "Bayesian Dynamic Programming." *Advances in Applied Probability* 7, 330-348.

Ross, S. (1983). "Introduction to Stochastic Dynamic Programming." New York, Academic Press.

Rothschild, M. (1974). "A Two-Armed Bandit Theory of Market Pricing." 9, 185-202.

Royden, H. L. (1988). "Real Analysis." New York, Macmillan.

Schwartz, L. (1965). "On Bayes' Procedures." 4, 10-26.

Strasser, H. (1985). "Mathematical Theory of Statistics." Berlin, Springer-Verlag.

Whittle, P. (1982). "Optimization Over Time: Dynamic Programming and Stochastic Control (I)." New York, Wiley.